DOCUMENT RESUME

ED 384 640 TM 023 808

AUTHOR Gallagher, Jo D.
TITLE A Multi-Disciplinary Approach to Integrating Evaluation and Training.
PUB DATE Apr 95
NOTE 11p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adult Education; Business Education; *Evaluation Methods; Evaluation Utilization; Human Resources; Information Dissemination; *Innovation; *Integrated Activities; *Interdisciplinary Approach; Labor Force Development; Models; *Program Evaluation; Quality Control; Total Quality Management; *Training
IDENTIFIERS *Diffusion of Innovations Research

ABSTRACT

A problem that exists in business and industrial training in particular, and in adult education in general, is that program evaluation is treated as an add-on activity, usually performed only at the request of someone outside the training organization. The relevant knowledge in two separate fields, quality improvement and innovation diffusion, can be merged with training program evaluation to suggest specific activities in which human resources development change might engage to integrate evaluation and training. An organization concentrating on quality improvement should incorporate customer focus, total involvement, assessment, systematic support, and continuous improvement into its endeavors. The study of the innovation diffusion has indicated that an individual's adoption or rejection of an innovation takes place in the individual through five stages on the way to adoption. This knowledge can also be built into the evaluation, regardless of the evaluation model chosen. (Contains 15 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Jo D. GALLAGHER
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

ED 384 640

A Multi-Disciplinary Approach to Integrating Evaluation and Training

Jo D. Gallagher
Florida International University
Miami, Florida 33199

Paper Presented at the Annual Meeting of
American Educational Research Association
San Francisco, California, April 1995

TM023808

BEST COPY AVAILABLE

Abstract: How can evaluation be fully integrated into the programs and services offered by HRD departments as standard operating procedure? This author takes the position that the relevant knowledge in two separate fields—quality improvement and innovation diffusion—can be merged with training program evaluation to suggest specific activities in which HRD change agents might engage to achieve this goal.

The purpose of this paper is to help human resource development practitioners solve a problem that exists in business and industrial training in particular, and in adult education in general; that of the treatment of program evaluation as an add-on activity, usually performed only at the request of someone outside the training organization. In many cases, the only evaluation data gathered stem from the "smile sheets" filled out by training participants at the end of a course that capture, at best, immediate reactions to the experience. This problem is not uncommon, and is not confined to evaluation. Training needs assessment appears to suffer the same treatment (Rossett, 1991).

If the general systematic training design process (which includes analysis, design, development, implementation, and evaluation) is compared with that of the continuous improvement process advanced by Shewhart (1931), better known as the *Plan*, *Do*, *Check*, *Act* (PDCA) cycle, several parallels can be discerned. For example, analysis and design are planning activities and evaluation is a checking activity. In considering the implementation of both processes in American industry in general, much emphasis has traditionally been on the action-taking parts (development and implementation; *Do* and *Act*) without much concern for planning and checking. It is no wonder then why training evaluation, embedded in the business and industrial context, is viewed as it is.

The consequences of this approach to evaluating training can be several. Consider the following scenarios:

1. Training programs in an organization are not evaluated at all, and unless there is a total breakdown in a classroom session which must be dealt with at the time it happens, little is done to improve training otherwise. The changes that are made in-session are not noted and not incorporated into further offerings of the course.
2. Data are collected by the training department but are not used to enhance programming. In some cases data are reported to an entity outside the training department (e.g., vice president of human resources) that may not do anything other than collect figures.
3. Evaluation, when it is performed, is done so as an afterthought, and less thoroughly than if planned along with the training it is to inform. Full benefits are, therefore, not realized for the costs involved.
4. Evaluation is performed in that data are collected and analyzed to discern necessary modifications which should be made to programs. The emphasis in the training department is heavily on the side of development, however, and few, if any, of the indicated changes are made. If anyone asks if programs are evaluated, staff can point to the data and reports of intended changes.

The adoption of total quality management (TQM) by an organization has changed this situation for the better in some cases. As some organizations have embraced TQM, some of the former operational emphasis on *Do* has spread to the other three phases and *Action* for improvement is taken based on data generated in the *Check* stage. It has been this author's experience, however, that the tremendous increase in training that occurs as a company adopts TQM tends to put even more emphasis on *Do* for the internal training unit because the demand for training comes from all over the company, not just from the incumbents in one job or even in a cluster of jobs. The press is to generate (buy, modify, create new) training that will enable all employees to function competently to implement (*Do*) the principles, practices, and tools of quality improvement. In some cases, this produces a situation similar to scenario number one, above. If this happens in organizations that follow the teachings of the quality gurus, what about the organizations that have not chosen to do so? They would appear to have even less of an incentive to evaluate their training programs. The situation is curious, given that most instructional design models include evaluation in one form or another. The problem is probably not, therefore, a lack of training on the part of the instructional development group.

If the lack of integration of evaluation and training is approached as a human performance problem, Leshin, Pollock and Reigeluth (1993) would suggest that if training in evaluation skills and knowledge is, indeed, not the solution, then the options include the following: (a) consider the incentive system, (b) consider making organizational changes, and (c) install a system to provide informational feedback. How then can program evaluation be built into the structure of the services a training unit provides to its constituent learners? The answer proposed in this paper was fashioned by combining knowledge gained from the research in two areas—quality improvement and communication, specifically, innovation diffusion—with training program evaluation. The relevant aspects of quality improvement and innovation diffusion are presented first, followed by a description of the components that form an integrated model with training program evaluation. For each component, its implications for practice are then presented.

Quality Improvement

Built on a foundation of its culture and values, a quality organization contains some key elements. Various authors have proposed their favorites, but there are some that appear consistently across the literature (Organizational Dynamics, 1988; Lewis & Smith, 1993). These key elements are the driving forces behind the various quality improvement movements regardless of their initials. A quality training organization would be well served to ensure that all its efforts, evaluation included, incorporate the following:

- Customer focus
- Total involvement
- Assessment
- Systematic support
- Continuous improvement

Customer focus refers to identifying and meeting the requirements and expectations of internal customers (other departments, employees, etc. within the company) as well as external customers (the ultimate recipient of the organization's product or service). The exchange of services, information, and so forth links internal and external customers together in a customer-supplier chain. This key element includes using a systematic approach to identify and meet the requirements of internal and external customers in such a way as to align what is produced with what is needed.

Total involvement means that everyone in the organization takes the responsibility for achieving quality; that it is not left to a select few or, perhaps worse, that no one is held accountable. It also means that all levels of the organization from top management to individual contributors are involved in achieving quality as it relates to what they specifically produce. This implies a respect for people closest to the work to make decisions that will ultimately insure customer satisfaction.

Assessment involves monitoring progress made toward quality goals—taking baseline measures and keeping track. What gets assessed (called a quality indicator) is heavily related to customer requirements. Who does the assessing should be closest to the work being done to meet those customer requirements. This also includes the use of facts and data as the primary basis for decision-making and as input for further enhancement of services and programs.

Within any organization, the policies, standards, and decisions support efforts to one degree or another. To the extent that these follow a quality strategy, quality services will prevail. This *systematic support* also refers to recognizing and rewarding individuals who use a quality strategy to achieve goals and removing barriers that prevent or dampen the use of quality strategies and tactics.

In a total quality environment, *continuous improvement* becomes the prevailing mind-set; a continual striving to improve that which is produced, whether that is a trained employee, a strategy for recruiting employees, or even a staff meeting. The idea is to look inside the organization and elsewhere for benchmarks—approaches or processes used to meet customer requirements that may be better than what is currently done. In this case, data gathered regarding quality indicators are analyzed for clues to improving the services to which they are related. This is done on an ongoing basis (e.g., through the PDCA cycle) as is communicating the successes of improvement efforts across the organization.

Innovation Diffusion

If evaluation of training can be viewed as a novel idea within the organization, or, if not a novel idea, one that has yet to gain wide adoption on a long-term basis, then the research on the diffusion of innovations can be brought to bear on integrating evaluation into the structure of training services and programs. Due to the limited scope of this paper, only the most salient points from the research will be made here.

Those who wish to forge this integration are change agents, and, as such, need to be prepared for the journey. Among other things, astute change agents know the following:

1. The process of innovation adoption, and that of change in general.
2. What and who adopters and opinion leaders are and how they facilitate further adoption.
3. How to use communication channels to foster adoption.
4. The various roles of the change agent.

Simply put, the adoption or rejection of an innovation takes place within the mind of an individual. The individual passes through five stages on the way to adoption, which can be thought of as full and ongoing use of an idea. Individuals within an organization may be at any one of these stages regarding the adoption of the idea of evaluation as part of the fabric from which training services and programs are fashioned. The stages in this adoption process (Rogers & Shoemaker, 1971) are as follows:

1. *Awareness*—the individual learns of the existence of a new idea but has no information about it.
2. *Interest*—the individual develops an interest in the new idea and searches for more information about it.
3. *Evaluation*—the individual mentally applies the new idea to his or her present and anticipated future situation and decides whether or not to try it; the "Is there anything in this for me?" question.
4. *Trial*—the individual actually uses the idea on a small scale to determine its usefulness.
5. *Adoption*—the individual uses the new idea full scale and ongoing.

As with any innovation, there are those who adopt it early—within the time span between first glimmer of an idea to full and widespread implementation—and those who adopt it later, or in some cases, not at all. The earlier adopters, if they can be identified, are useful to the change agent. They are not necessarily any different than later adopters in age, but they tend to be less dogmatic and possess a greater ability to deal with abstraction. Within this earlier adopter category change agents will find a higher degree of opinion leadership and more social participation. Although not all early adopters are also opinion leaders, those that are can facilitate the adoption of an innovation by others because they are accessible and respected. People can and do talk to them, as well as listen to what they have to say.

A change agent's success is linked to many things, some of which are within his or her control. A short primer on change agentry can provide some insight. First, a change agent's success is linked to working through opinion leaders. A change agent working alone will not succeed. As opinion leaders are also many times early adopters, they can be enlisted to carry momentum and persuade later adopters. Second, a change agent must use two channels of communication—mass media and interpersonal—for full adoption of an innovation. Mass media channels such as memos, flyers, newsletters, and formal written or oral announcements at large gatherings are good for promoting awareness and changes in weakly held attitudes. Interpersonal channels, which involve face-to-face interchange between people, can help overcome some of the social and psychological barriers to change as well as persuade individuals to form or revise more

strongly held attitudes. Third, an individual's satisfaction with an adoption decision is positively related to his or her level of participation in the decision-making. A savvy change agent will provide varied and numerous opportunities for people to get involved.

An Integrated Model and Its Implications

In order to take advantage of what is known from the research in quality improvement and innovation diffusion, this knowledge must be built into evaluation efforts. Fortunately, this can be done whether evaluation efforts are yet to be planned or exist in some manner currently. The first ingredient, however, is a solid approach to training program evaluation. Slick change agentry aside, if the evaluation scheme is not a good one, few people will be convinced to use it; those that become convinced may not remain so for long.

The purpose of this paper is *not* to present yet another evaluation model, but to present components that, when considered as part of evaluation, facilitate the integration of evaluation into the structure of the services and programs that human resource development (HRD) departments provide. These components can be overlaid onto an evaluation model of the evaluator's creation or choosing and are designed to be used as tactics in the long-term strategy to foster adoption of the model to full and widespread use. There are already many good program evaluation models from which to choose. Popham (1993) has classified some of the major ones into the following categories: Goal-attainment models, judgmental models, decision-facilitation models, and naturalistic models. Whether the model selected is sophisticated and customized (Rendulic, 1994) or as straightforward as that of Craig (1978), it should be an articulated and agreed-upon plan for what will be evaluated, how it will be evaluated, to what audience(s) it will be reported, and how the information is to be used.

The approach proposed here is flexible enough to accommodate most any evaluation model and adds components favorable to structural integration with the system of training services to be evaluated. As the figure on the last page shows, the evaluation model is key. Four components surround and enhance the selected model based on the research. Each component will be discussed in turn.

Top-Down Support and Recognition

Time spent in program evaluation activities is time spent *not* doing something else. There must be some payoff or incentive for those engaged in evaluation. Although there may be some who derive intrinsic joy from this type of activity, more of us do things because we are externally rewarded for it in some way. This refers back to Leshin, Pollock and Reigeluth's (1993) first suggestion and to stage three of the innovation adoption process where the prospective adopter asks "what's in this for me?" If the answer to this question is "nothing" or "not much," then it is unlikely that program evaluation will be integrated. Top-down support and recognition can and should take many forms. Administrators, for example can drive fear out of evaluation by doing such things as the following: removing obstacles to data sources; encouraging the use of evaluation data for program and service enhancement; requesting that budget dollars be justified, in part, at least, by evaluation results; arranging release time for those heavily engaged in evaluation activities; instituting a policy of program review that requires decisions based on data rather than on managerial power plays. Recognition ranges from

a simple pat on the back for a job well done to testimonial banquets and beyond. The key is to get support and recognition from high up in the organization and from the levels in between to the people doing the work. Depending on the culture of the HRD organization, recognition from peers can be a powerful reinforcer as well. This component, when present in the planning, delivery and evaluation of training can remove some of the social and psychological barriers to adoption and reflects the systematic support principle of quality.

Participation from Critical Mass of Stakeholders

In any evaluation effort there are many tasks to be done: planning, gathering and analyzing data, reporting, and decision-making to name just a few. If training programs and services are to be evaluated, then the people providing and managing these should be involved in that evaluation, perhaps with the technical assistance from others where necessary. This may sound like the fox guarding the henhouse, but it reflects the second key element of quality, total involvement, and fosters a mind-set of continuous improvement, the fifth key element. Also, when people in close proximity are working on some of the same things, interpersonal communication channels are used to exchange ideas. Earlier adopters can help later adopter types to acquire an interest in evaluation and assess its utility for themselves.

Participation does not mean that everyone does everything. This is why the component specifies a "critical mass." Critical mass is achieved when the major opinion leaders are involved and some of the skeptical have been convinced. At that point the effort begins to take on a momentum of its own. Trying out a selected evaluation model on a single training program gets a small number of people involved while providing feedback about enhancing not only the instructional program but also the evaluation process itself. From this trial run, participant evaluators can speak from experience about the process as they facilitate others to do it. Where possible, opinion leaders should be involved as data sources as well as participant evaluators to maximize impact on interpersonal channels of communication.

The first step in gaining participation is to determine what distinct tasks need to be performed. This may be as simple as looking closely at the evaluation process selected. The next step is to determine who is the best person for the task. "Best" can be defined as (a) can do the task, (b) is respected by peers, and (c) communicates well with others inside and outside the evaluation effort. It may be that the best "person" is actually a group. The third step is to start recruiting and getting people ready for the task(s) they are to do. It also helps to have several names ready in case the first choice declines or is otherwise engaged. One way to offer opportunities for involvement is to first form an evaluation model selection team. Once the model is chosen, another team can be formed to pilot it on one program where those associated with the program (e.g., designers and trainers) are willing. A third team can assess the utility of the evaluation model and modify it where necessary to permit wider use. The membership on these three teams may overlap to provide continuity. These tactics to gain participation may also be viewed as organizational changes to enhance human performance (Leshin, Pollock & Reigeluth, 1993).

The members of the teams should be prepared to participate. Change agents can provide the necessary training or briefing to ensure successful participation. For example, change agents may provide information about candidate models of evaluation such as features, expected impact, and how they work. They also may assist participant evaluators with creating data collection forms and interpreting the resulting data.

Multi-Referenced Interpretation of Evaluation Data

Many models for program evaluation are based on comparing actual performance with intended outcomes or goals. Their lineage dates back to Ralph Tyler and the Eight Year Study of the 1930s (Tyler, 1942). To put this into quality language, an organization would first determine its internal and external customers' requirements, which then would become its intended outcomes or objectives, and then measure its performance against them. This criterion-referenced approach to interpreting evaluation results is fitting and proper, as is norm-referenced interpretation, known in quality circles as "competitive benchmarking" whereby organization "A" examines organization "B" which offers a similar product or service and has a good reputation so that organization "A" can make its own operation better. For example, the training department from a transportation company may examine the training department in a nearby fast food corporation to gain insight as to how the transportation group could improve its distribution of and registration for non-classroom-based courses across the country. Data collected on the fast food training group would then be compared to that of the transportation group to give a norm-referenced view of how well it was doing compared to another training unit that had successfully faced a problem common to both units. Some may argue that if the customers' requirements are satisfied, there is no need to engage in competitive benchmarking. This view is limited, however. There may be other organizations that can reveal ways to save money, for example, while maintaining a high level of service.

A variation is to compare current performance against baseline data, collected at some earlier time. The baseline data provide the benchmark for comparison on critical quality indicators so that progress can be tracked. Of course, if no baseline data exist, they must be collected first.

Data-Based Continuous Improvement of Training and the Evaluation System

Michael Scriven gave the field of education the term *formative* evaluation (Scriven, 1967) to describe a focus of data-gathering designs and analysis techniques on the improvement of instructional sequences that were still under development. Scriven contrasted this approach with that of *summative* evaluation, which focused on determining the worth of completed instructional programs and sequences. The formative approach, proactive in nature, uses evaluation data to enhance in some way an already well-planned effort. This is similar to what is advocated in this component of the integrated model. This does not negate the utility of summative evaluation, but does point out the need for data-based decisions in pursuit of continuous improvement.

Both training and the evaluation system (model, process, etc.) are cited in the heading of this component section because both are important beneficiaries of evaluation data. Too often, however, the field of focus for evaluation and improvement efforts stops short of the evaluation system itself (meta-evaluation). As evaluation provides feedback

to the rest of the overall system, it is part of that system and should be treated as such. This means that data must be collected about the evaluation efforts and those data used to enhance those same efforts. Following Fullam and Ponfret (1977), meta-evaluation during implementation of evaluation should focus on facilitating implementation and local system capabilities, rather than on judging the success or failure of the chosen model. For example, following a pilot evaluation of a sequence of sales training programs, a simple debriefing of the people involved revealed that some of the intended data sources were not used because few systematic records were kept by supervisors. This information became input for figuring out how to track past trainees so that future evaluation data-gathering efforts would be easier without overburdening supervisors. By using evaluation data to inform meta-evaluation, adoption of evaluation as an innovation is facilitated for those who were trial users (the people who had trouble tracking down past trainees) in the fourth stage of the adoption process and for those who remained skeptical until someone else field tested it. The model benefits from data-driven enhancements, thus making it more attractive to potential adopters who are in the second (interest) or third (evaluation) stage of adoption. Another way to approach the collection of meta-evaluation data is to hold a series of "what worked/didn't work" sessions both during and after evaluating a program or service. The members of the evaluation team thus become part of a focus group for enhancing not only their future efforts, but also the efforts of others.

As an element of systematic support, program reviews (cycles of program implementation, evaluation, and enhancement) can be scheduled so that data are gathered over a fixed period of time and are used as the basis for subsequent program enhancements. Consider a ten-day technical training course that is offered quarterly (February, April, June, and October). Evaluation data are collected over the year's term as the program is implemented and trainees go back to their jobs. Those data are then analyzed and interpreted following the October class and indicated changes are made before the February offering. Where multiple courses are offered, review cycles can be staggered so that the workload is balanced among course delivery, evaluation data gathering, and program revision. Where courses are offered more often, sampling procedures may be used for gathering data across the delivery term.

References

Craig, D. P. (1978). *Hip pocket guide to planning and evaluation.* Austin, TX: Learning Concepts.

Deming, W. E. (1982). *Out of the crisis.* Cambridge, MA: MIT Center for Advanced Engineering Study.

Dick, W. and Carey, L. (1990). *The systematic design of instruction* (3rd ed.). Harper Collins Publishers.

Fullam, M. & Ponfret, A. (1977). Research on curriculum and instruction implementation. *Review of Educational Research 47*(2), 335-97.

Gay, L. R. (1985). *Educational evaluation and measurement: Competencies for analysis and application* (2nd ed.). Columbus, OH: Charles E. Merrill.

Leshin, Pollock, & Reigeluth, C. (1993). *Instructional Design strategies and tactics.* Englewood Cliffs, NJ: Educational Technology Publications.

Organizational Dynamics, Inc. (1988). *The concepts of quality.* Boston, MA: Author.

Popham, W. J. (1993). *Educational evaluation.* Boston, MA: Allyn and Bacon.

Rendulic, P. (1994). *An investigation into the potential value and the validity of a curriculum evaluation model for community colleges.* Unpublished doctoral dissertation, Florida International University, Miami.

Rogers, E. M. and Shoemaker, F. F (1971). *Communication of innovations: A cross-cultural approach* (2nd ed.). New York: The Free Press.

Rossett, A. (1991). Needs assessment. In G.J. Anglin (Ed.). *Instructional technology: Past, present, and future.* Englewood, CO: Libraries Unlimited, Inc.

Scriven, M. (1967). The methodology of evaluation. In R. E. Stake (Ed.). *American Educational Research Association Monograph Series on Evaluation, No.1.* Chicago: Rand McNally.

Shewhart, W. A. (1931). *Economic control of quality of manufactured product.* New York: Van Nostrand.

Tyler, R. W. (1942). General statement on evaluation. *Journal of Educational Research, 35,* 492-501.

Worthen, B. R. and Sanders, J. R. (1973). *Educational evaluation: theory and practice.* Worthington, OH: Charles A. Jones Publishing.